Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 132 (2017) 705–710

Energy Procedia

www.elsevier.com/locate/procedia

11th Nordic Symposium on Building Physics, NSB2017, 11-14 June 2017, Trondheim, Norway

Quasi-Monte-Carlo-based probabilistic assessment of wall heat loss

Tianfeng Hou[a]*, Dirk Nuyens[b], Staf Roels[a], Hans Janssen[a]

[a]KU Leuven, Department of Civil Engineering, Building Physics Section, Kasteelpark Arenberg 40 – box 2447, BE-3001 Heverlee, Belgium
[b]KU Leuven, Department of Computer Science, Numerical Analysis and Applied Mathematics section, Celestijnenlaan 200a - box 2402, BE-3001 Heverlee, Belgium

Abstract

In this paper, the potential of quasi-Monte Carlo methods for uncertainty propagation is assessed, via a case study of heat loss through a massive masonry wall. Four quasi-Monte Carlo sampling strategies – Optimized Latin hypercube, Sobol sequence, Niederreiter-Xing sequence and Good Lattice sequence – are applied and compared. Moreover, in order to terminate the quasi-Monte Carlo simulation when the desired accuracy is reached, an error estimation method is implemented. The outcomes show that all the four quasi-Monte Carlo methods outperform the standard Monte Carlo method; the Niederreiter-Xing sequence and Sobol sequence tend to be the best.

Peer-review under responsibility of the organizing committee of the 11th Nordic Symposium on Building Physics.

Keywords: Quasi-Monte Carlo; uncertainty propagation; building performance; sampling efficiency; convergence monitoring

1. Introduction

Monte Carlo based probabilistic analysis is widely used in many uncertainty assessments, because of its general applicability and typical robustness. However, in cases where the deterministic core simulation is relatively time consuming, the number of needed repetitions of that core simulation should be minimized. In that respect, three criteria are important: 1) the sampling scheme should be as efficient as possible, 2) monitoring of the sampling convergence

* Corresponding author. Tel.: +32 16328832; fax: + 32 16321980
E-mail address: tianfeng.hou@kuleuven.be

Peer-review under responsibility of the organizing committee of the 11th Nordic Symposium on Building Physics
10.1016/j.egypro.2017.10.010

should be possible, and 3) sequential additions of single samples should be possible in order to stop the Monte Carlo when a sufficient accuracy is obtained.

The current state of the art in building physics in that respect is replicated optimized Latin hypercube sampling [1], wherein smaller subdesigns are sequentially added to reach the desired number of runs n (instead of using a single n-run optimized Latin hypercube design). This is not wholly in line with the third criterion, as blocks of samples are sequentially added instead of single samples. What is more, the main drawback of this method is that correlations may exist between the smaller subdesigns, given that each small subdesign is optimized with the same algorithm. And these correlations may in turn lead to a bias for the error estimation, as well as a decline of the sampling convergence rate. Instead of the replicated optimized Latin hypercube sampling, in this paper we will use lattice rules and digital nets to construct quasi-Monte Carlo sampling schemes for uncertainty propagation, exemplified through a case study of the transmission heat loss through a massive masonry wall.

Below, first the sampling strategies are briefly put forward, with focus on potential problems that they may bring in. Subsequently, the calculation object and its input parameters are introduced, as that forms the central application in this study. The following two sections respectively analyse the sampling efficiency and error quantification for these sequential sampling methods, which constitute the crucial issues of this investigation. Finally, conclusions on which sampling scheme is considered most optimal are formulated.

Nomenclature

n:	the number of samples	$\overline{Q_{n,r}}(f)$:	mean of r n-run QMC evaluations
d:	the number of parameters	r:	the number of replications
μ:	mean of cumulated heat loss	θ:	standard error
σ:	standard deviation of cumulated heat loss	$Q_n(f)$:	a single n-run QMC evaluation
n^{QMC}:	the number of samples of one QMC sequence	QMC:	quasi-Monte Carlo
n^{MC}:	the number of samples of one MC sequence	MC:	Monte Carlo

2. Sampling strategies

In this paper, a Monte Carlo based uncertainty analysis is used for propagating the randomness of the input to the randomness of the output. Traditionally, the standard Monte Carlo uses pseudo-random sequences of size n to quantify the probability distribution of the target outcomes, which typically provide a convergence rate of $O(n^{-1/2})$ [2]. In contrast, quasi-random sequences are constructed with quasi-Monte Carlo methods, which can offer a better uniformity than the random sequence, resultantly covering the domain of interest quickly and evenly, and thus yielding a convergence rate of $O((\log n)^d \cdot n^{-1})$ for continuous functions [2]. As mentioned before, there are two main families of quasi-Monte Carlo methods: lattice rules and digital nets which are constructed by different approaches to achieve uniformity of the points. More details of lattice rules and digital net can be found in [2]. Before we move to using quasi-Monte Carlo methods for building physical applications, it is important to mention two crucial challenges of quasi-Monte Carlo, as well as to explain how to overcome these difficulties.

First, contrary to standard Monte Carlo which uses independent random samples to estimate the probability distribution of the target outcomes, quasi-Monte Carlo methods apply deterministic quasi-random sequences. Thus, the conventional error estimation method based on the central limit theory is not valid for quasi-Monte Carlo. Secondly, in order to let quasi-Monte Carlo outperform standard Monte Carlo with respect to sampling efficiency, $O((\log n)^d \cdot n^{-1})$ needs to be smaller than $O(n^{-1/2})$. Therefore, d should be relatively small and n should be relatively large [2].

To overcome the first challenge, we can make use of randomization techniques to obtain multiple independent quasi-random sequences, and in this way the confidence interval can be constructed based on their standard error. A survey of different randomization techniques for quasi-Monte Carlo integration can be found in [2]. With respect to the second challenge, Van Gelder [3] shows that in many building physical engineering problems, even though the number of the total input parameters may be large, the number of effective parameters is usually fairly small. And based on Wang [4], for low effective dimension problems, the randomized quasi-Monte Carlo sequences can perform much better than standard Monte Carlo, even for large d.

In addition to optimized Latin hypercube designs, three other quasi-random sequences (Sobol [5], Niederreiter-Xing [6] and Good Lattice [7]) are applied and compared. Here, the Sobol and Niederreiter-Xing sequences are examples of digital nets, while the Good Lattice sequence is constructed by lattice rules. For randomization, the Sobol sequence applies a random linear scramble combined with a random digital shift [2]. On the other hand, the Good Lattice and Niederreiter-Xing sequences are randomized by Cranley-Patterson shifting and digital shifting, respectively [2]. Moreover, all three low-discrepancy sequences are in base 2, and the Good Lattice sequence has been optimized for an unanchored Sobolev space, and the optimization is for up to 2^{13} points [8].

3. Calculation object and input parameters

3.1 Calculation object

Today, 30% of the European building stock consists of 'historic' buildings built prior to World War II. These buildings are typically far less energy-efficient than new buildings, and they hence account for a large share of the total energy consumption of buildings. One important option to reduce their energy consumption is to install internal insulation, if external insulation cannot be considered for aesthetical reasons. Internal insulation is however often associated with moisture damage, and much care should be taken when applying this solution. This paper is part of the EU H2020 RIBuild project [9], which aims at developing effective and comprehensive guidelines for internal insulation in historic buildings. Given that many uncertainties are to be considered in those guidelines, RIBuild has adopted a probabilistic approach, and hence this article on the use of quasi-Monte Carlo for uncertainty propagation.
The reference situation prior to retrofit is commonly a massive masonry wall, and that configuration is adopted here as calculation object. In order to judge the feasibility of internal insulation in historic buildings, the hygrothermal performances of internally insulated massive walls – heat loss, mould growth, wood rot, frost spalling, … – need to be investigated [10]. To reduce the calculation load in this study, this paper limits that performance assessment to the transmission heat losses through the wall. The target outputs are the mean and the standard deviation of the distribution of cumulated heat loss over one year.
To do so, the thermal behaviour of the wall is simulated with Delphin, wherein the conductive heat transfer equation is solved under the relevant interior and exterior boundary conditions. The exterior boundary conditions include convection as well as short and long wave radiation, directed by the climate data of one of three German cities (see Table 1). At the interior surface, combined convection and long wave radiation is imposed, governed by the interior air temperature as describe in EN 15026 [11].

3.2 Probabilistic and deterministic input parameters

The thermal analysis involves a number of parameters, related to the geometry, materials, climates as well as surface properties. We separate all input parameters into two categories, deterministic and probabilistic, which respectively have fixed values and probability distributions, see Table 1.

4. Sampling efficiency

The sampling efficiency can be described as the number of repetitions of the core deterministic simulation needed to reach a desired accuracy. More specifically, the deviation from reference solution, the variance or root-mean-square error of the targeted outcomes are often used as indicators of the accuracy. In other words, sampling strategies requiring less runs of simulations are termed 'more efficient'. In this paper, we assess the sampling efficiency of the standard Monte Carlo, uniformity-based optimized Latin Hypercube, Sobol, Niederreiter-Xing and Good Lattice, by comparing the deviations of their outcomes from a reference solution. That reference solution is obtained via 10 replications of a 2^{15}-run randomized good lattice sequence and its generating vector is optimized for up to 2^{20} points [8], amounting to 194.0888 kWh and 45.3854 kWh for respectively the mean and the standard deviation of the resulting heat loss distribution. To assess the accuracy of the reference solution, the standard errors of these 10 replications are calculated, which is 0.0017 for the mean and 0.0105 for the standard deviation, respectively.

Table 1. Deterministic and probabilistic input parameters.

Input parameters	Input distributions		
Climatic conditions			
Meteorological climate data	D(Essen, Bremerhaven, Munchen)		
Exterior surface			
Convective heat transfer coefficient (W/m²K)	Essen N(7.34, 2.11) Bremerhaven N(11.67, 3.17) Munchen N(7.22, 1.77)		
Solar absorption coefficient (-)	U(0.4, 0.8)		
Long-wave emissivity (-)	0.9		
Wall orientation (degree from North °)	U(0, 360)		
Brick layer			
Thickness (m)	U(0.15, 0.50)		
Material	D(Brick 1, Brick 2, Brick 3) (see below)		
Interior surface			
Total heat transfer coefficient (W/m²K)	8		
Explanation of symbols used:			
• *U(a,b): uniform distribution between a and b;*			
• *D(a,b): discrete uniform distribution with options a and b;*			
• *N(μ, σ): Normal distribution with mean μ and standard deviation σ;*			
Material property	Brick 1	Brick 2	Brick 3
Bulk density (kg/m³)	2087	1786	1980
Thermal capacity (J/kgK)	870	1000	834
Dry thermal conductivity (W/mK)	0.9	1.08	0.996

4.1 Simulation results

As mentioned in section 2, contrary to standard Monte Carlo, the evaluation of a single quasi-Monte Carlo sequence yields a deterministic result, and there is thus no standard error output on the estimator. Thus, multiple quasi-Monte Carlo sequences need to be created and the mean of these independent quasi-Monte Carlo evaluations

$$\overline{Q_{n,r}}(f) = \frac{1}{r}\sum_{i=1}^{r} Q_n^{(i)}(f) \tag{1}$$

is taken as the final approximation. Hence, the total number of function evaluations for quasi-Monte Carlo is $r{\cdot}n^{QMC}$. In this paper, we take r = 10 and $n^{QMC} \in [2^3,\ldots,2^7]$. In order to make a fair comparison with the standard Monte Carlo sampling, we should therefore take the number of samples in the standard Monte Carlo method $r^{MC} = r{\cdot}n^{QMC}$, and hence $n^{MC} \in 10{\cdot}[2^3,\ldots,2^7]$. In addition, for obtaining a robust conclusion of the convergence behavior of each sampling scheme, except for uniformity-based optimized Latin Hypercube, 10 independent replications of $r{\cdot}n^{QMC}$ quasi Monte Carlo evaluations are made at each of these 5 sample sizes. Since the generation cost for optimized Latin Hypercube designs is very high, no replications are generated for that particular scheme. The sampling efficiency of standard Monte Carlo, uniformity-based optimized Latin Hypercube, Sobol, Niederreiter-Xing and Good Lattice, for the mean and standard deviation of the cumulated heat loss distribution, are shown in Figure 1.

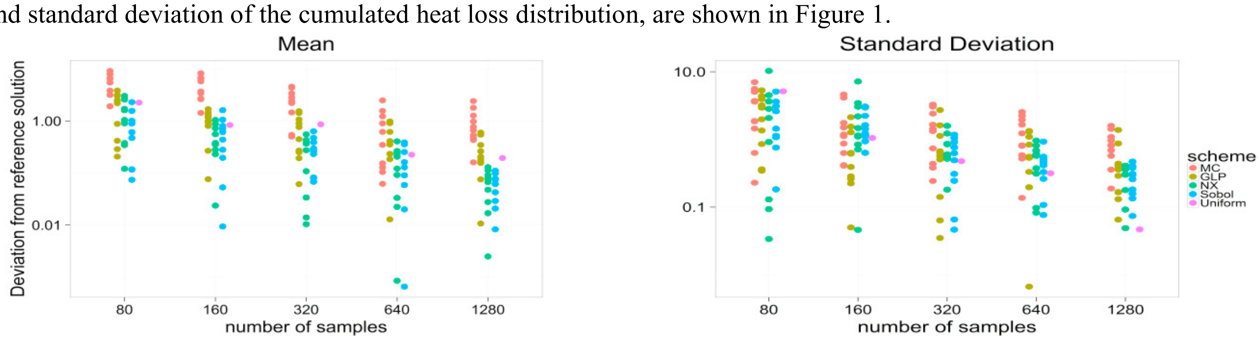

Figure 1: the sampling efficiency of standard Monte Carlo (MC), uniformity-based optimized Latin hypercube (Uniform), Sobol (Sobol), Niederreiter-Xing (NX) and Good Lattice (GLP), with respect to the mean and standard deviation of the cumulated heat loss

4.2 Discussion

In Figure 1, it can be seen that all the quasi-Monte Carlo sampling strategies outperform the standard Monte Carlo method for both the mean and standard deviation. It is also clear that the uniformity-based optimized Latin Hypercube scheme has a similar sampling efficiency as the other three quasi-Monte Carlo sequences. It is however less versatile than the digital nets and lattices, and should thus not be preferred. First, generating optimized Latin Hypercube designs is very time-consuming: their generation cost grows exponentially with the size of the sampling design [1]. Additionally, the number of samples needs to be determined before starting the Monte Carlo evaluations and they do hence not support sequential sampling. Third, as there is no randomization technique for optimized Latin Hypercube design, it is difficult to obtain independent optimized Latin Hypercube designs for error estimation.

In addition, we can also see that the convergence rate of Good Lattice sampling is slightly slower than Sobol and Niederreiter-Xing: this is most probably due to the fact that the generating vector applied in this paper has not been optimised for our specific calculation object. Hence, to potentially improve the performance of the Good Lattice sequence, a generating vector aligned with the particular type of the calculation object needs to be developed. In summary, the use of Sobol sequence and Niederreiter-Xing sequence are currently preferred.

5. Error estimation

The randomized quasi-Monte Carlo method takes advantage of quasi-Monte Carlo, which yields both a higher convergence rate, while at the same time making error estimation possible. The target is to use the standard error of several independent randomized quasi-Monte Carlo sequences to construct the confidence interval around the outputs [2]. More concretely, to estimate the error of the quasi-Monte Carlo analysis, r independent randomized quasi-Monte Carlo sequences are evaluated. These can be obtained by adding a random or digital shift $\Delta \in [0,1]^d$ to the initial quasi-Monte Carlo sequence and repeat the procedure r times, where $\Delta = (\Delta_1,\ldots,\Delta_d)$. Hence, each of the repetitions is an independent quasi-Monte Carlo sequence (which has low-discrepancy properties) and therefore the standard error of these independent approximations can be derived:

$$\theta = \sqrt{E|I(f) - \overline{Q_{n,r}}(f)|^2} \approx \sqrt{\frac{1}{r(r-1)}\sum_{i=1}^{r}(Q_n^{(i)}(f) - \overline{Q_{n,r}}(f))^2} \tag{2}$$

Here, the expectation is taken with respect to the random shifts. Next, based on the Chebyshev's inequality, a 75% confidence interval can be obtained by using $\overline{Q_{n,r}}(f) \pm 2\,\theta$. In this paper, we simply take $2\,\theta$ as our error indicator. In order to evaluate the accuracy and convergence speed of our error indicator, the ratio of the error indicators of each sampling schemes to the deviation of the respective quasi-Monte Carlo approximation from the reference solution, as well as the convergence behaviour of these error indicators for the mean of the cumulated heat loss, are shown in Figure 2.

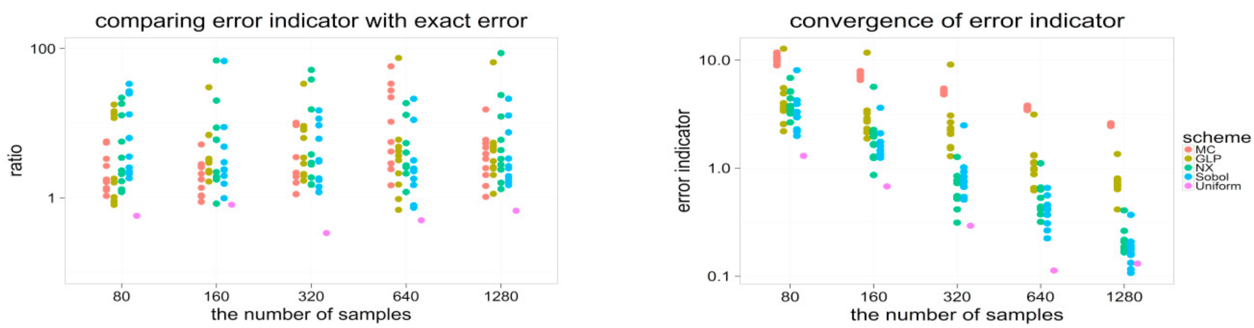

Figure 2: the ratio of the error indicators of each sampling schemes to the deviation of the respective quasi-Monte Carlo approximation from the reference solution (left) and the convergence behaviour of error indicators (right), with respect to the mean of the cumulated heat loss distribution

It is shown in Figure 2 (left), except for the uniformity-based optimized Latin Hypercube scheme, that the ratio of the error indicators of the other three sampling schemes (Sobol, Niederreiter-Xing and Good Lattice sequence) to the deviations of their respective quasi-Monte Carlo approximations from the reference solution are almost always bigger than 1 and often smaller than 10, which implies that their error indicators are always larger than their exact errors and usually within the difference of one magnitude. Hence, the error of the three quasi-Monte Carlo methods can be quantified accurately and conservatively based on the error estimation method. On the other hand, for the uniformity-based optimized Latin Hypercube scheme, their error indicators are however always smaller than their exact errors. The reason has been mentioned before: because there is no randomization technique for optimized Latin Hypercube design, correlations may exist between the smaller subdesigns and thus, an error indicator derived by the standard error of these smaller subdesigns may lead to a too optimistic error estimation.

In addition, it is shown in Figure 2 (right) that the error indicator of all the sampling strategies decrease as the number of the evaluations increase, with respect to the mean of the cumulated heat loss, and it is also shown that this indicator decreases faster for quasi-Monte Carlo methods than standard Monte Carlo method. Moreover, comparing between the Sobol, Niederreiter-Xing and Good Lattice sequences, it is also demonstrated that this indicator declines faster for the Sobol and Niederreiter-Xing sequences than for the Good Lattice sequence. All in all, in combination with the conclusion of section 4.2, the use of Sobol sequence and Niederreiter-Xing sequence are currently favoured, based on their sampling efficiency and error estimation.

6. Conclusion

In this report, we studied both the sampling efficiency and the error estimation methods for different quasi-Monte Carlo sampling strategies, based on a case study of heat loss through a massive masonry wall. It is shown that the Sobol sequence and Niederreiter-Xing sequence have the highest sampling efficiency and the error of quasi-Monte Carlo methods can be estimated accurately and conservatively based on the standard errors which are constructed by implementing the related randomization techniques.

References

[1] Janssen, Hans. "Monte-Carlo based uncertainty analysis: Sampling efficiency and sampling convergence." Reliability Engineering & System Safety 109 (2013): 123-132.

[2] L'Ecuyer, Pierre, and Christiane Lemieux. "Recent advances in randomized quasi-Monte Carlo methods." Modeling uncertainty. Springer US, 2005. 419-474.

[3] Van Gelder, Liesje. "A Probabilistic Design Methodology for Building Performance Optimisation An Application to Low-Energy Dwellings." (2014).

[4] Wang, Xiaoqun, and Kai-Tai Fang. "The effective dimension and quasi-Monte Carlo integration." Journal of Complexity 19.2 (2003): 101-124.

[5] Bratley, P., and B. L. Fox. "Algorithm 659 Implementing Sobol's Quasirandom Sequence Generator." ACM Transactions on Mathematical Software. Vol. 14, No. 1, 1988, pp. 88–100.

[6] Pirsic, Gottlieb. "A software implementation of Niederreiter-Xing sequences." Monte Carlo and Quasi-Monte Carlo Methods 2000. Springer Berlin Heidelberg, 2002. 434-445.

[7] Nuyens, Dirk. "The construction of good lattice rules and polynomial lattice rules." (2014): 223-256.

[8] https://people.cs.kuleuven.be/~dirk.nuyens/qmc-generators/

[9] http://ribuild.eu/

[10] Vereecken, Evy, et al. "Interior insulation for wall retrofitting–A probabilistic analysis of energy savings and hygrothermal risks." Energy and Buildings 89 (2015): 231-24

[11] Hygrothermal performance of building components and building elements. Assessment of moisture transfer by numerical simulation.